

## RESEARCH ARTICLE

# Power smoothing in large wind farms using optimal control of rotating kinetic energy reserves

Simon De Rijcke,<sup>1</sup> Johan Driesen,<sup>1</sup> Johan Meyers<sup>2</sup><sup>1</sup> Department of Electrical Engineering/ Electa, KU Leuven, Kasteelpark Arenberg 10 box 2445, 3001 Heverlee, Belgium.<sup>2</sup> Department of Mechanical Engineering, KU Leuven, Celestijnenlaan 300A, 3001 Leuven, Belgium.

## ABSTRACT

In large wind farms, self-induced turbulence levels significantly increase the variability of generated power in a range of time scales from a few seconds to several minutes. In the current study, we investigate the potential for reducing this type of variability by dynamically controlling the rotating kinetic energy reserves that are present in the farm's wind turbines. To this end, we reduce the burden of frequency regulation on remaining conventional units when they are displaced in favor of wind turbines. We focus on the development of a theoretical benchmark framework in which we explore the trade-off between high energy extraction and low variability using optimal coordinated control of multiple turbines subject to a turbulent wind field. This wind field is obtained from a large-eddy simulation of a fully developed wind farm boundary layer. The controls that are optimized are the electric torque and the pitch angles of the individual turbines as function of time, so that turbines are accelerated or decelerated to optimally extract or store energy in the turbines' rotating inertia. Results are presented in terms of Pareto fronts (i.e. curves with optimal trade-offs), and we find that power variations can be significantly reduced with limited loss of extracted energy. For a one-turbine case, such an optimal control leads to large potential reductions of variability, but mainly for time scales below 10 sec if we limit power losses to a few percent. Variability over longer time scales (10–100 sec) is reduced considerably more for coordinated control. For instance, restricting the energy-loss incurred with smoothing to 1%, and looking at time scales of 50 sec, we manage to reduce variability with a factor of 6 for a coordinated case with 24 turbines, compared to a factor of 1.4 for an uncoordinated case. Copyright © 2013 John Wiley & Sons, Ltd.

## KEYWORDS

wind farm; turbulence; power variations; optimal turbine control

## Correspondence

Simon De Rijcke, Department of Electrical Engineering, research group Electa, KU Leuven, Kasteelpark Arenberg 10 – box 2445, 3001 Heverlee, Belgium.

E-mail: [simon.derijcke@esat.kuleuven.be](mailto:simon.derijcke@esat.kuleuven.be)

Received ...

## 1. INTRODUCTION

A growing number of incentives for the use of renewable energy sources has given the wind power market an incredible boost during the past decade. As a result, farm owners have shifted their interest from small wind farms equipped with small units, to larger offshore wind farms with a turbine capacity up to several Megawatts. The average capacity of offshore wind farms grew from around 80 MW in 2007 to 300 MW for farms currently under commission [1]. However, a larger share of wind power in the generation mix forces grid operators to rely on wind power to maintain a stable operating power system. Today, primary reserves or Frequency Containment Reserves (FCR) stabilize frequency fluctuations, and secondary reserves or Frequency Restoration Reserves (FRR) restore the frequency to a reference, i.e. 50 or 60 Hz [2]. The fast FCR are primarily guaranteed by the action of governors, coupled to conventional power plants. In addition, the inertia from synchronous generators introduces a time delay that gives the governors time to react to frequency variations. With more synchronous generators replaced by wind power plants, the equivalent system inertia and available governing action are reduced (unless frequency response is required from all generating units), but also more power variability is added to the

power system. Although power reserves are primarily designed to cover severe frequency events, they are used more often during normal operation. As a result many system operators would favor the capability of wind farms to limit ramp rates and power gradients. Although such limiting is not possible in many countries because of legislation, this is already the case in Ireland and Denmark [3, 4]. In the future, regulations may become stricter, e.g., by penalizing large power gradients, or market mechanisms may change, e.g., by compensating for operating reserves. Under these circumstances, it may become economically beneficial to actively smooth the power output of wind farms. In the current work, we investigate possible trade-offs between power smoothing and energy extraction in wind farms, focusing on turbulence-induced fluctuations. To this end, we study optimal coordinated control of electrical torques and pitch angles of variable-speed wind turbines to smooth power fluctuations in wind farms that result from atmospheric turbulence. We aim at optimally employing the rotating kinetic energy reserves of the turbines in the wind farm. This allows us to reduce power variability for time scales ranging from one second to a few minutes. The mechanisms leading to these reductions are carefully explained further in this paper. Our approach presumes variable-speed wind turbines. Time constants that relate to the detailed electrical topology of these machines (for instance a doubly-fed induction generator or full converter topology) are at least one order of magnitude smaller, and of no further importance in the current study.

In wind farms, changing weather patterns, wake effects, atmospheric turbulence, etc., create a time varying wind field, leading to large power variations and gradients over a wide range of time scales. Turbulence-induced variability is mainly prevalent in a range of time scales from seconds up to 15 minutes, directly impacting fast wind farm power gradients, and thus influencing power regulation and not load following [5]. In large wind farms, self-induced turbulence levels are much higher compared to the case of lone-standing turbines [6], further increasing the problem of turbulence-induced power variability in modern farms. Today, striking unbalance problems are avoided by a limited share of wind in a flexible generation mix, consisting of fast-acting gas turbines. However, this balancing comes at a cost that gas turbine owners want to be reimbursed for. Although characteristics of wind power fluctuations are the subject of earlier studies [7, 8, 9, 10] (mostly with the aim to determine the characteristics and accordingly the required reserves), the ability of large-scale wind farms to smooth power variations remains largely unexplored.

In the past, first attempts to smooth power have focussed on single turbine control [11, 12, 13], without including interaction between individual turbines. More recently, some studies have looked into the coordination within a wind farm [14, 15], mainly focussing on the control architectures, but testing them for simple artificially generated wind variations. In the current study, we focus on power smoothing in large wind farms with fluctuating velocity fields that are obtained from time-resolved turbulent-flow simulations, i.e. based on large-eddy simulations [16, 17]. This provides realistic turbulent wind data with correct temporal and spatial coherence, allowing us to study effects of geometrical placement of controlled turbines (e.g. streamwise versus spanwise arrangement). We further focus on optimal control of rotating kinetic energy reserves in the wind farm. Although optimal control is difficult to implement in reality, it provides a benchmark for optimal trade-offs between power smoothing and energy extraction in wind farms.

Wind-turbine operation is often classified into three regions: I–III [18]. The first region is at very low wind speeds where aerodynamic forces cannot overcome the turbine's internal friction losses. At very high wind speeds (Region III), the power output of turbines is restricted by loading constraints on its mechanical structures and by economical constraints on the size of the power generator. In this region, turbine power is controlled at a constant level, independent of wind speed, and thus, power variability remains limited. In Region II, power output is not restricted, and wind turbines adapt their rotational speed and power output to the wind speed, yielding large variability in power output. Consequently, we focus in the current work on Region II operation, and do not impose limits on the rotor speed.

The paper is further organized as follows. First, §2 elaborates on the optimal control methodology and problem set-up. Secondly, §3 explains the optimal control results for a single turbine, with a focus on the interpretation of the energy cost that comes with smoothing power variations. Thereafter, the benefit of a coordinated control approach with multiple turbines is highlighted in §4. Finally, §5 summarizes the conclusions.

## 2. OPTIMAL CONTROL: METHODOLOGY AND SET-UP

First, §2.1 introduces the formulation of the optimization problem with objective functions and constraints. Next, §2.2 elaborates on the model for the aerodynamic torques, and the wind series used in this study. Finally, §2.3 discusses the computational set-up used to perform the simulations to obtain the results in §3 and §4.

### 2.1. Formulation of the optimization problem

We consider the optimal control of a wind farm by means of the electrical torques  $\mathbf{T}_e(t)$  and pitch angles  $\beta(t)$  of all turbines in the farm. Here,  $\mathbf{T}_e$  and  $\beta$  are vectors containing respectively the torques  $T_{e,i}(t)$  and pitch angle  $\beta_i(t)$  of individual turbines, with  $i = 1 \dots N$  turbines. Both control variables dynamically depend on time and are optimized in a time window  $0 \leq t \leq T$ . By using these control variables, we aim at two competing objectives: i.e. maximization of total

energy extraction  $\mathcal{J}_1$ , and minimization of power fluctuations  $\mathcal{J}_2$ , where:

$$\mathcal{J}_1(\omega(\beta, T_e), T_e, T) \equiv \int_0^T J_1(\omega, T_e) dt = \int_0^T -P_f(\omega(t), T_e(t)) dt = \int_0^T -\sum_{i=1}^N T_{e,i}(t) \omega_i(t) dt, \quad (1)$$

$$\mathcal{J}_2(\omega(\beta, T_e), T_e, T) \equiv \int_0^T J_2(\omega, T_e) dt = \int_0^T \left( P_f(\omega(t), T_e(t)) - P_f(\omega(t-\tau), T_e(t-\tau)) \right)^2 dt, \quad (2)$$

with  $P_f$  the total farm power, and  $\omega(\beta, T_e, t)$  the vector with rotational speeds of the different turbines. Following standard conventions in optimization, we choose the signs of  $\mathcal{J}_1$  and  $\mathcal{J}_2$  such that the optimization is formulated as a minimization problem. Optimization of torques  $T_{e,i}(t)$  and pitch angles  $\beta_i(t)$  to the first objective  $\mathcal{J}_1$ , maximizes the energy yield of the entire wind farm, simultaneously coordinating all  $N$  turbines' operating points. The second objective  $\mathcal{J}_2$  minimizes power variability, by penalizing power output differences at time instances that are separated by a user-selected time lag  $\tau$ . Since  $P(t) - P(t-\tau) = \int_{t-\tau}^t (dP/dt') dt'$ , the cost function  $J_2$  is equivalent to penalizing power gradients that are averaged over a time window  $\tau$ . In the current work, we choose  $\tau = 1$  sec (cfr. §2.3 for further discussion).

The cost functionals  $\mathcal{J}_1$  and  $\mathcal{J}_2$  constitute competing objectives, i.e. it is easily seen that a trivial optimum for  $\mathcal{J}_2$  is reached for  $P_f = 0$  (no variability, but also no power). Therefore, we formulate an optimization problem based on a weighted sum of both cost functionals, i.e.

$$\mathcal{J} = \alpha \mathcal{J}_1 + (1 - \alpha) \mathcal{J}_2 \quad (3)$$

$$= \int_0^T (\alpha J_1(\omega, T_e) + (1 - \alpha) J_2(\omega, T_e)) \cdot dt \equiv \int_0^T J(\omega, T_e) \cdot dt \quad (4)$$

where we vary  $\alpha \in [0, 1]$  in order to construct Pareto fronts (cfr. §3 and §4), and explore different trade-offs between  $\mathcal{J}_1$  and  $\mathcal{J}_2$ .

The farm power  $P_f$  appearing in  $\mathcal{J}_1$ , and  $\mathcal{J}_2$  depends on the controls  $T_e$  but further also on the rotational speeds  $\omega$  of the turbines. The latter is a state variable that is simply determined by Newton's second law for rotation, i.e. for turbine  $i$ :

$$I \frac{d\omega_i}{dt} = T_{a,i}(V_{\infty,i}, \omega_i, \beta_i) - T_{e,i}. \quad (5)$$

where  $T_{a,i}$  is the aerodynamic torque, and  $I$  the total moment of inertia. The aerodynamic torque depends on the wind speed  $V_{i,\infty}$  at turbine hub height (for turbine  $i$ ), which we obtain from large-eddy simulations. Further details, and a model for  $T_{a,i}$  are provided in §2.3. Finally, the optimal control problem that is considered then corresponds to

$$\min_{\omega, T_e} \int_0^T J(\omega, T_e) \cdot dt, \quad (6)$$

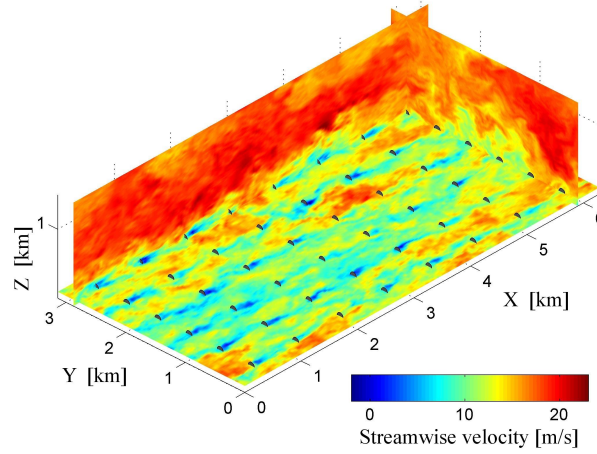
$$\text{s.t.} \quad I \frac{d\omega}{dt} - T_a(V_{\infty}, \omega, \beta) + T_e = 0, \quad (7)$$

## 2.2. Wind field and aerodynamic torque model

In the current section, we discuss the determination of the aerodynamic torque in Eq. (5). To that end, detailed velocity time series are obtained from large-eddy simulations of a very large wind farm in a fully developed neutral planetary boundary layer [16, 17]. In particular, the large-eddy simulations are performed in a physical domain containing  $6 \times 8$  turbines, separated by  $7.85D$  in streamwise direction, and  $5.24D$  in spanwise direction, with  $D$  the turbine rotor diameter. The simulations use periodic boundary conditions in horizontal planes, representing fully developed boundary-layer conditions. Such a case is relevant to wind farms that are 10–20 times longer than the typical height of the atmospheric boundary layer (1 km).

Turbines are represented by actuator-disk models, in which the turbine-thrust force acting on the flow is proportional to the local turbine-disk velocity  $V_{d,i}$ , and a local thrust coefficient  $C'_T$  representing the accumulated effect of lift and drag coefficients of the turbine blades (cfr. Meyers & Meneveau [17] for details). A typical snapshot of the turbulent incident wind field in the wind farm is shown in Figure 1. For the current study, thirty thousand successive snapshots are generated every 1 sec, and turbine-disk velocities  $V_{d,i}$  are retrieved. Sampling every 1 sec is sufficient as wind speed variations up to a few seconds are filtered by the rotor disk.

Changes in angular velocity of the turbines, and control of turbine pitch angles directly impact on  $C'_T$  as a result of changing angles of attack, and lift and drag coefficients of the blade. In principle, such changes would couple back to the turbulent flow, since an increase of force locally slows down the flow, and vice versa. However, it is not in the scope of the current work to solve an optimal control problem in which the flow equations are fully coupled, as this is beyond



**Figure 1.** Typical snapshot of the turbulent streamwise velocity field in a wind farm boundary layer simulation [16, 17].

current computational resources (cfr. Ref. [19] for the computational challenges related to one flow-coupled wind farm optimization). Instead, we presume the turbulent flow field (obtained from the LES) to be insensitive to our control actions. Such a one-way coupled approach is justified provided that variations in  $C'_T$  remain small. **This is indeed** the case if the total extracted energy for  $\alpha < 1$  (cfr. Eq. (3)) remains close to the maximum at  $\alpha = 1$ , such that the average value of  $C'_T$  cannot change **too much** from its optimal value. Remark that there is no simple relation between the value of  $\alpha < 1$ , and the related energy loss, as this also depends on the optimization case (i.e. number of turbines, their placement, etc.). In §3, and §4, we will construct Pareto fronts for the full range  $\alpha \in [0, 1]$ , including regions where energy losses are significant. In these regions the Pareto fronts give a qualitative appreciation of power smoothing only. However, for discussions and conclusions, we focus on power smoothing for which the incurred energy loss remains limited to a few percentages only, i.e. regions where the Pareto fronts are expected to be accurate.

Thus, in the current work,  $C'_T$  is kept at an optimal value in the LES simulations [17], and we only employ them to obtain turbine disk velocities  $V_{d,i}$ , but no aerodynamic torques. Instead, in order to formulate the effect of control actions on the aerodynamic torque, we use a simple heuristic fit of the turbine power coefficient  $C_p$  to calculate the turbine torque, i.e. for turbine  $i$ :

$$P_i \equiv \frac{1}{2} C_{p,i} \rho V_{i,\infty}^3 A = T_{a,i} \omega_i \Rightarrow T_{a,i} = \frac{C_{p,i} \rho V_{i,\infty}^3 A}{2 \omega_i}, \quad (8)$$

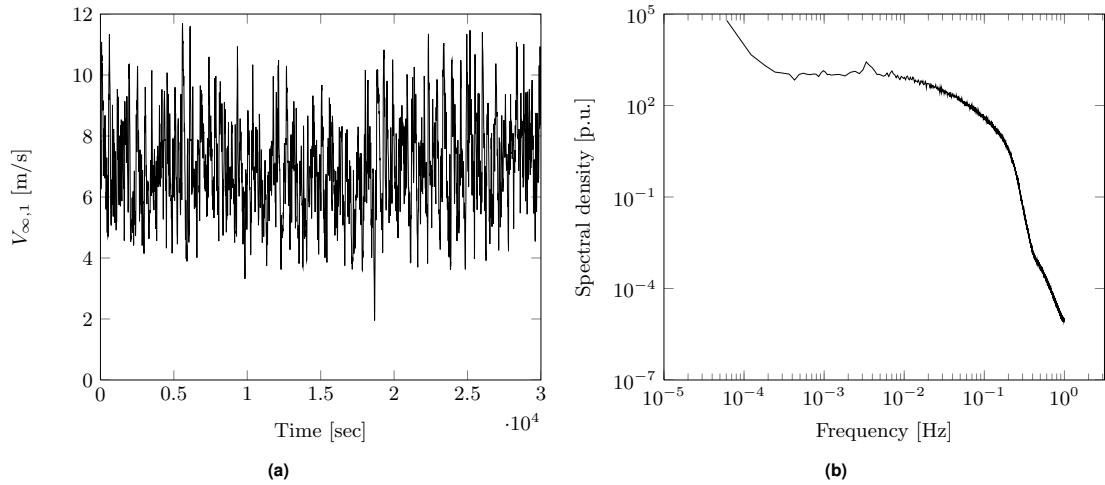
with  $A = \pi D^2/4$  the rotor area. We use an empiric relation for  $C_p$  that corresponds to [20, 21]

$$\begin{aligned} C_p(\lambda_i, \beta_i) &= c_1 (c_2 \gamma_i + c_3 \beta_i + c_4) e^{c_5 \gamma_i}, \\ \text{with } \gamma_i &= \frac{1}{\lambda_i + c_6 \beta_i} + \frac{c_7}{\beta_i^3 + 1}, \end{aligned} \quad (9)$$

and where  $\lambda_i = R \omega_i / V_{i,\infty}$  is the tip-speed ratio. Further,  $c_1 = 2.20 \times 10^{-1}$ ,  $c_2 = 1.16 \times 10^2$ ,  $c_3 = -4.00 \times 10^{-1}$ ,  $c_4 = -5.00$ ,  $c_5 = -1.25 \times 10^1$ ,  $c_6 = 8.00 \times 10^{-2}$  and  $c_7 = -3.50 \times 10^{-2}$ .

The model for  $C_p$  in Eq. (9) is calibrated for a wind-turbine in lone-standing conditions, and depends on the ‘unperturbed’ upstream wind velocity  $V_{i,\infty}$ . This velocity differs from the velocity at disk level  $V_{d,i}$  that is lower than  $V_{i,\infty}$ , since the flow is locally decelerated by the turbine. However, using the calibration conditions for  $C_p$ , we can obtain  $V_{i,\infty}$  from  $V_{d,i}$  using classical 1D momentum theory. This yields  $V_{i,\infty} = V_{d,i} (1 + C'_T/4)$ . For the case used here,  $C'_T = 4/3$  (cfr. case A3 in Calaf et al [16]), leading to  $V_{i,\infty} = V_{d,i} 4/3$ . This leads to a time series  $V_{i,\infty}$  ( $i = 1 \dots N$ ) that is ranged in region II of turbine operating, i.e. the operating region where optimization of power extraction typically leads to large variability. Note that the exact rescaling of  $V_{i,\infty}$  as function of  $V_{d,i}$  is further not important, as we will never look at absolute power output in the result discussion, but rather always normalize with respect to an uncontrolled reference (so that the scaling factor drops out of the result).

Finally, in Figure 2a we show the time-series  $V_{i,\infty}$  for one turbine, illustrating the large temporal variability of the velocity in wind farms. In Figure 2b, the spectral density of the wind signal is given, averaged over the 48 turbines in the simulation.



**Figure 2.** Wind speed series of one turbine (a) and the spectral density of the wind signal averaged over 48 turbines (b).

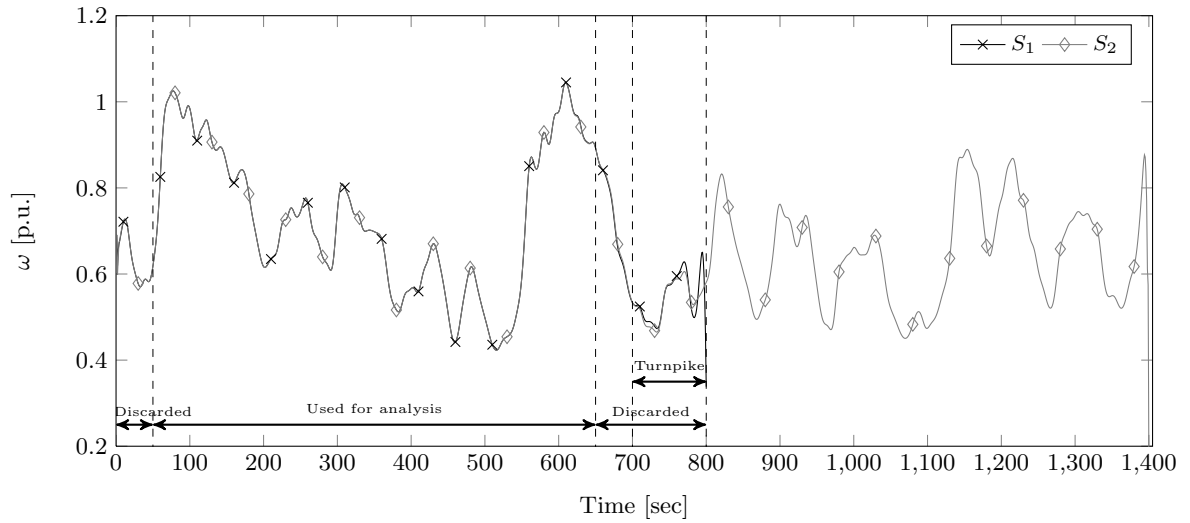
### 2.3. Computational set-up

To solve the optimization problem, the system is first discretized in time. To that end, the state variables  $\omega(t)$  are sampled every 0.2 sec, and Eq. (7) is discretized using the trapezoid rule according to Newton–Cotes. The controls  $\beta(t)$  and  $T_e(t)$  are allowed to change every 1.0 sec and remain constant during the intermediate time steps of the states. Given the frequency content of the wind signal ( $< 0.5\text{Hz}$ ), and size of the inertia of the turbines, we verified that this discretization gives sufficient accuracy, i.e. further time-step refinement does not lead to noticeable differences in results.

The optimization problem stated in Eqs. (6–7) is solved by a gradient based solver from TOMLAB/KNITRO [22]. This solver offers an iterative Conjugate Gradient (CG) approach for large scale problems, that is combined with a Sequential Linear-Quadratic Programming (SLQP) optimizer for fast convergence of results close to the optimum. The gradient of the cost-function is obtained by automatic differentiation that provides an automatic formulation of the discrete adjoint system, and is facilitated by the matrix formulation of the problem in the TOMLAB environment. An analytic formulation of the gradient calculation is also elaborated in Appendix I by formulating the continuous adjoint equations [23]. We verified that the TOMLAB gradient provides the same results as an evaluation based on a discretization of the continuous adjoint equations. For the different optimization cases in the current work, we use the automatic differentiation from TOMLAB as it provides a very efficient and flexible formulation that allows for easy adjustment of objective functions, etc.

For the time-lag  $\tau$  in Eq. (2), we select 1 sec. Thus, fast variability is reduced first. We found that selecting too large values for  $\tau$  (e.g. 60 sec) penalizes only power differences that are separated by 60 sec. This choice leads to poor results, with sudden jumps in power output, and large fluctuations at smaller time scales. Low values of  $\tau$  (in the range of 1–10 sec) lead to similar results. Moreover, we also find that the more effective power smoothing becomes (when performing coordinated control on larger wind farms), the more variations are also damped at larger time scales (cfr. detailed discussion in §4).

For the optimization time window, we select  $T = 800$  sec. However, as a result of the nature of the optimization problem, all turbines decelerate to standstill at the end of the optimization time horizon (cfr. Figure 3 for a one-turbine optimization case). In fact, since the optimization is not considered beyond  $t = T$ , any remaining rotational energy in the turbines is considered useless by the algorithm, and instead it is fully depleted in favor of a larger total energy yield. As is appreciated in Figure 3, this only occurs at the very end of the time horizon, i.e. towards  $t = 800$  sec for  $S_1$ . This phenomenon is also known as the turnpike problem [24]. To avoid any effects of this unrealistic deceleration of turbines, we evaluate optimized results only up to  $t = 650$  sec, discarding the last 150 sec of the optimization time window. In addition, non-optimal initial conditions result in transients that decay very fast in time at the beginning of each simulation run. As a consequence, we also discard the first 50 sec of the optimization window. We further verified that elongating the time horizon does not impact the results. In Figure 3 an optimization result ( $S_2$ ) for  $T = 1400$  sec is compared with  $S_1$  for  $T = 800$  sec, showing that results do not differ within time window  $t \in [50, 650]$  sec that is used for analysis.



**Figure 3.** Rotor speed for two simulations. The first series  $S_1$  of  $T = 800$  sec runs over  $t \in [0, 800]$  sec. The second series  $S_2$  of  $T = 1400$  sec runs over  $t \in [0, 1400]$  sec. Time series of 800 sec are used for simulating, of which only the results within  $t \in [50, 650]$  sec are used for analysis.

### 3. OPTIMAL CONTROL RESULTS FOR A SINGLE TURBINE

The optimal control of a single turbine is discussed. First, in §3.1 Pareto fronts with the trade-off between energy yield and power variability are introduced for a single turbine, and reductions in variability are further interpreted. Subsequently, we investigate optimal control of electric torque with or without optimal control of the blade pitch angles in §3.2. Finally, in §3.3 we discuss averaged Pareto fronts.

#### 3.1. Energy yield versus power variations for a single turbine

The construction of a Pareto front allows one to compare optimal trade-offs between the two competing objectives  $\mathcal{J}_1$  and  $\mathcal{J}_2$  introduced in Eqs. (1) and (2). We construct such a Pareto front by using the weighted sum method for multi-objective optimization [25]. This method selects a priori a vector of ten weights in the range  $\alpha \in [0, 1]$  to create a set of solutions to this bi-objective optimization problem. In the current work, we will plot these Pareto fronts in terms of

$$E_* = \mathcal{J}_1 / [\mathcal{J}_1]^{\alpha=1}, \text{ and} \quad (10)$$

$$R_*^{1/2} = (\mathcal{J}_2 / [\mathcal{J}_2]^{\alpha=1})^{1/2}, \quad (11)$$

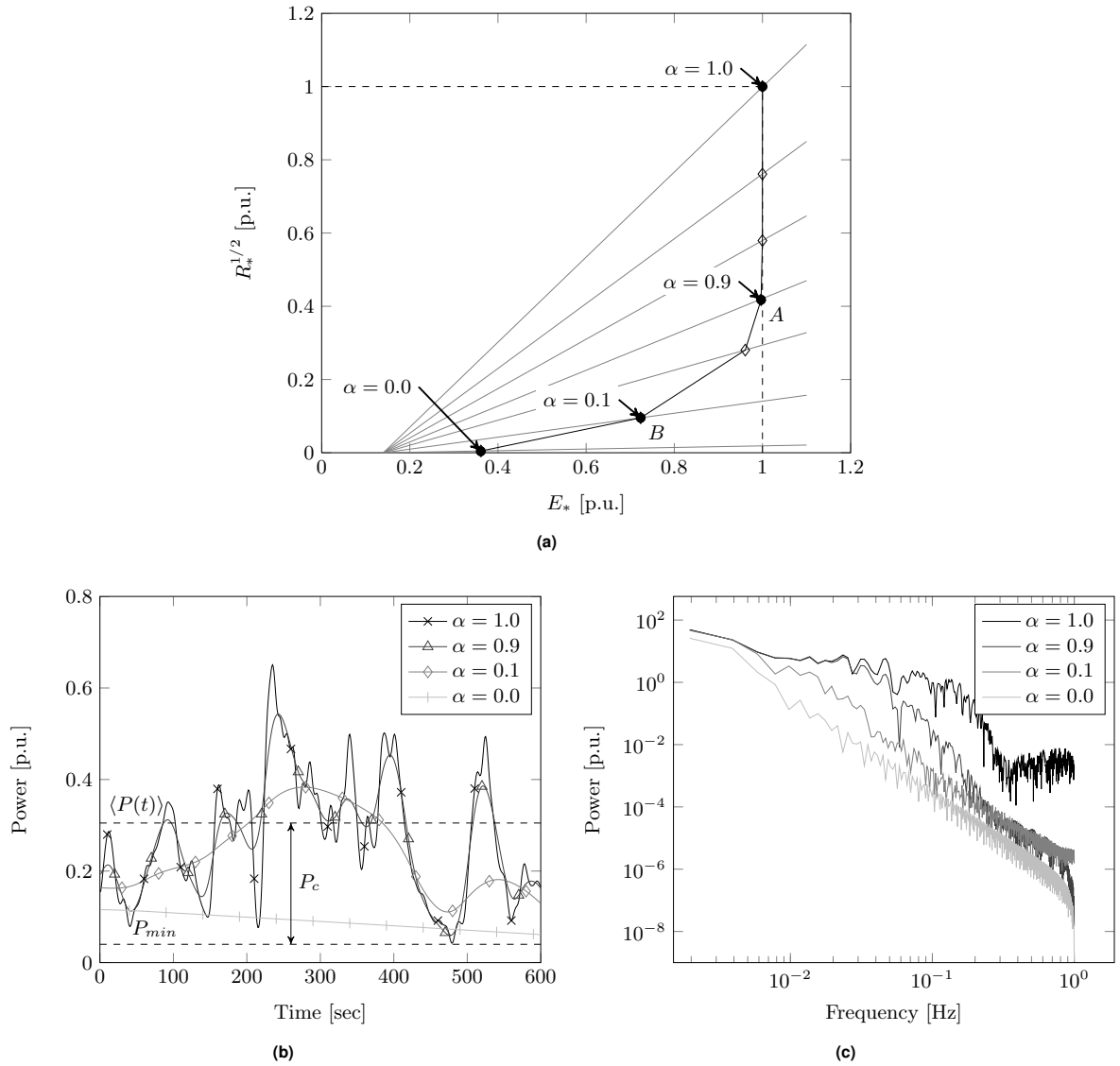
with  $E_*$  the energy extraction normalized with its maximum value at  $\alpha = 1$ , and  $R_*^{1/2}$  the square root of the normalized variability, which also has units of energy. Thus, for  $\alpha = 1$ , the Pareto fronts start at  $(1, 1)$ .

In Figure 4a a typical Pareto front is shown obtained for optimal control of one turbine. When decreasing  $\alpha$ , gradually penalizing variability in the multi-objective cost function  $\mathcal{J}$  (cfr. Eq. (3)), we observe initially that variability  $R_*^{1/2}$  decreases considerably at almost no loss of total energy yield (note that in this region, only fast fluctuations are effectively damped – cfr. further discussion below). Only if  $R_*^{1/2}$  is reduced below 0.4 [p.u.] the energy yield starts to drop, reaching a minimum value of 0.14 [p.u.] for  $R_*^{1/2} = 0$ .

In Figure 4b, we show the power extracted as function of time, and for different values of  $\alpha = 0.0, 0.1, 0.9, 1.0$  (also marked in Figure 4a). It is apparent that the electrical power signal is smoothed in time by lowering the weight  $\alpha$ . When lowering  $\alpha$  from 1.0 to 0.9, first power variations with high frequencies and smaller amplitudes are smoothed. This is also clearly visible in Figure 4c, in which the power density spectrum is displayed for the different cases. When further lowering  $\alpha$ , the reduction of power variations shifts to lower frequencies with a higher power density, and thus at a higher energy cost. To further interpret the relation between variability, and power yield, we can reformulate the cost function  $\mathcal{J}_2$  in terms of a time average, i.e.

$$\frac{1}{T} \mathcal{J}_2 = \frac{1}{T} \int_0^T (P(t) - P(t - \tau))^2 dt = \langle [P(t + \tau) - P(t)]^2 \rangle, \quad (12)$$





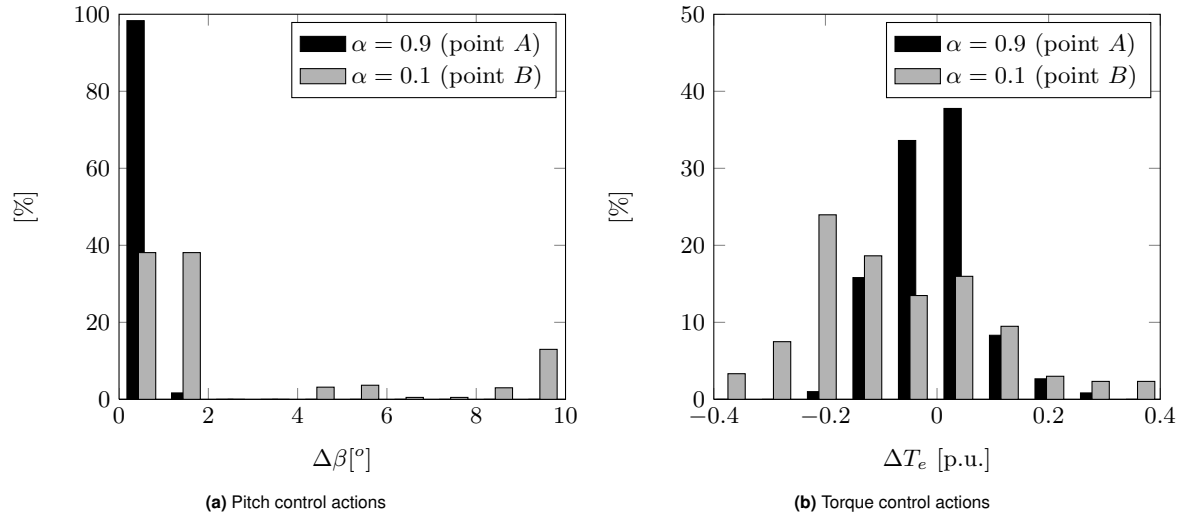
**Figure 4.** (a) Pareto front for a single turbine fed by a wind speed series of 600 sec. This front shows the optimal trade-off between a reduction of power variations and maximization of energy extraction, by varying the weight  $\alpha \in [0, 1]$ . Electrical power (b) and its mean power spectral density (c) for a single turbine after optimization with varying  $\alpha$ .

where we use the short-hand notation  $\langle \cdot \rangle$  for the time average. In order to further discuss this expression, and come to a useful interpretation of the optimal Pareto trade-offs, we first introduce

$$P(t) \equiv P_{min} + P_c[1 + p_v(t)], \quad \text{with } P_{min} = \min_{t \in [0, T]} P(t); \quad P_c = \langle P(t) \rangle - P_{min}, \quad (13)$$

and where by construction  $p_v(t) \equiv [P(t) - \langle P(t) \rangle] / P_c$  so that  $\langle p_v(t) \rangle = 0$ , and  $\min p_v(t) = -1$ . Thus, we split the signal in a constant minimum load  $P_{min}$ , a remaining mean average  $P_c$ , and fluctuating normalized power  $p_v$  (cfr. also Figure 4b, where  $P_{min}$ , and  $P_c$  are also plotted). Inserting in Eq. (12) yields

$$\begin{aligned} \langle [P(t + \tau) - P(t)]^2 \rangle &= P_c^2 \langle [p_v(t + \tau) - p_v(t)]^2 \rangle \\ &= P_c^2 (2 \langle p_v(t)^2 \rangle - 2 \langle p_v(t + \tau) p_v(t) \rangle) \\ &= 2 P_c^2 \langle p_v(t)^2 \rangle \left[ 1 - \frac{\langle p_v(t + \tau) p_v(t) \rangle}{\langle p_v(t)^2 \rangle} \right]. \end{aligned} \quad (14)$$



**Figure 5.** The number of pitch angle (a) and torque (b) deviations with respect to the optimal trajectory. The latter is obtained by only maximizing the energy yield, which corresponds to the Pareto point at  $\alpha = 1$  in Figure 4a.

In the last expression,  $\langle p_v(t)^2 \rangle$  corresponds to the variance of the fluctuating part, and  $\langle p_v(t + \tau)p_v(t) \rangle$  to its autocorrelation. Thus  $\mathcal{J}_2$  can be decreased by increasing the autocorrelation of the fluctuating normalized power, and decreasing its variance  $\langle p_v(t)^2 \rangle$  (remember that  $\min p_v(t) = -1$  is fixed by construction). Another way of decreasing  $\mathcal{J}_2$  is by decreasing  $P_c$ , but obviously, this is not beneficial, as it decreases the total extracted energy, i.e. remember that  $\langle P(t) \rangle = P_{min} + P_c$ , with a value  $P_{min}$  that is roughly fixed by the lowest wind speed in the time window  $T$ .

For the particular case that  $\mathcal{J}_2$  only decreases due to a reduction of  $P_c$ , we find that  $R_*^{1/2} \sim E_* - E_{*,min}$  (with  $E_{*,min} \equiv -P_{min}T/[\mathcal{J}_1]^{\alpha=1}$ ). This defines a set of lines through the point  $(E_*, R_*^{1/2}) = (0, E_{*,min})$ . In Figure 4a, these lines are also represented, and this is useful for further interpretation. Indeed, if we consider a point A on the Pareto fronts, the reduction of variability in this point in the direction  $\alpha \rightarrow 0$  is a result of ‘smart’ optimal control as long as the slope of the Pareto front in the point A is steeper than the slope of the line through A and  $(0, E_{*,min})$ . If this is not the case, further reductions of variability for  $\alpha \rightarrow 0$  are just achieved by reducing  $P_c$ .

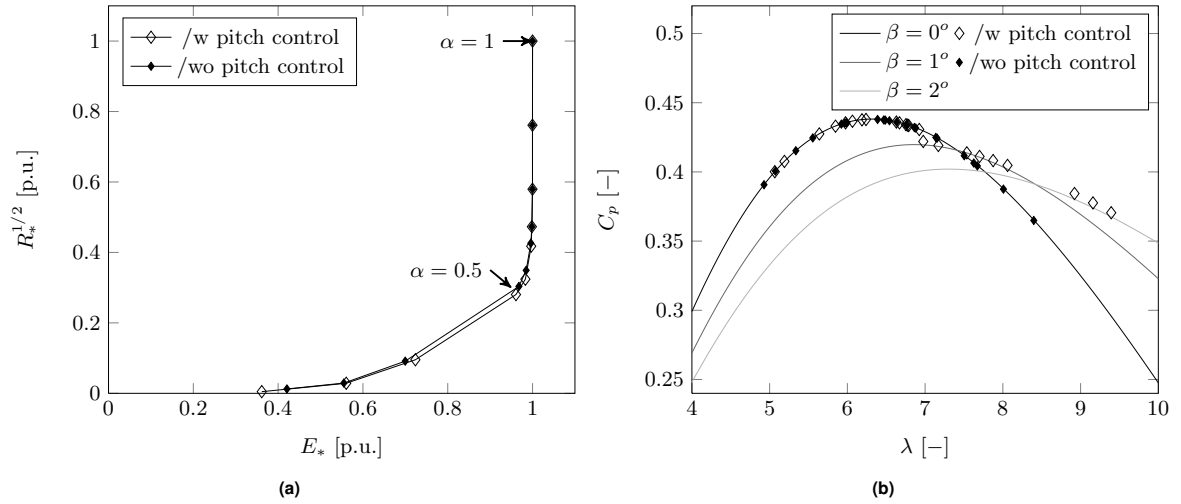
The analysis above demonstrates that the selected cost function  $J_2$  penalizes fluctuations around the mean power in an indirect manner. An obvious alternative may consist of directly penalizing the variance of the fluctuations. However, this would require a priori knowledge of the mean power  $\langle P(t) \rangle$ , which is not necessarily easily available in the context of a moving-horizon control algorithm.

### 3.2. Evaluation of optimized control actions

We now look in more detail at the optimized control actions used at different points on the Pareto front. Two points are selected, i.e. Point A and Point B, as marked on Figure 4a. The first point A corresponds to  $\alpha = 0.9$  and the second point point B to  $\alpha = 0.1$ . In Figure 5 we display histograms of the controls  $\beta$  and  $T_e$ . From these it is clear that in Point A, mainly the generator torque ( $T_e$ ) is used, while in Point B the number of pitch ( $\beta$ ) control actions increases considerably. Torque actions steer the kinetic energy exchange. This requires the rotor speed to deviate from the optimal tip-speed ratio, and causes the operating points to deviate from maximum efficiency (cfr. deviations from maximum  $C_p$  in Figure 6b). Such deviations reach values up to 13% above rated rotor speed for point B, possibly exceeding the manufacturers’ design tolerance. Such limits on rotor speed are not explicitly added here (cfr. also discussion in §5). Finally, when kinetic energy exchange does not suffice to achieve the level of power smoothing demanded by the optimizer, pitch actions are added. The latter directly reduce power extraction from wind. This explains the higher energy loss at point B with respect to point A.

In region II operation of single variable-speed turbines without power smoothing, pitch-control actions are normally not required. In this case, generator torque control is used to keep the turbine rotational speed and associated tip-speed ratio close to optimal, for which  $\beta = 0$  is further also optimal. However, when optimal power smoothing is used (as in point B), instantaneous tip-speed ratio’s will differ more-and-more from the optimal tip speed ratio, so that  $\beta = 0$  is not necessarily optimal anymore. In Figure 6b we show  $C_p$  curves for different pitch angles as function of  $\lambda$ . From this figure, it is clear





**Figure 6.** Comparison of results with and without pitch control. (a) Pareto fronts for a single turbine fed by two wind speed series A and B. (b) Coefficient of performance  $C_p$  for  $\alpha = 0.5$  compared to  $C_p$ -lines of fixed pitch angle  $\beta$ .

that for high values of  $\lambda$ , a pitch angle  $\beta > 0$  is optimal, and these values are effectively selected by the optimal control algorithm.

We further investigated optimal power smoothing using only generator torque control, and keeping  $\beta = 0$ . Results are shown in Figure 6a. We find that the Pareto front shifts slightly down when excluding pitch control actions, but differences are small. Thus, in practice power smoothing can also be realized using only generator torque control, which would avoid an increase in wear and tear on pitch systems.

### 3.3. Ensemble averaging of optimal results

Up till now, we discussed optimal power smoothing (of a single wind turbine) evaluated over a time window of 600 sec, and subject to a turbulent velocity field obtained from large-eddy simulations. However, turbulence in a wind-turbine boundary layer is characterized by large variability, and time scales up to several hundreds of seconds. Thus one single time series of 600 sec is in itself not sufficient to provide a statistically converged expectation of the gains of optimal power smoothing. To that end, the optimization results need to be averaged over a large set of different statistically independent turbulent time series. Therefore, we construct averaged Pareto fronts by using the averages

$$\bar{E}_* = \frac{\frac{1}{n} \sum_{j=1}^n \mathcal{J}_{1,j}}{\left[ \frac{1}{n} \sum_{j=1}^n \mathcal{J}_{1,j} \right]^{\alpha=1}}, \text{ and} \quad (15)$$

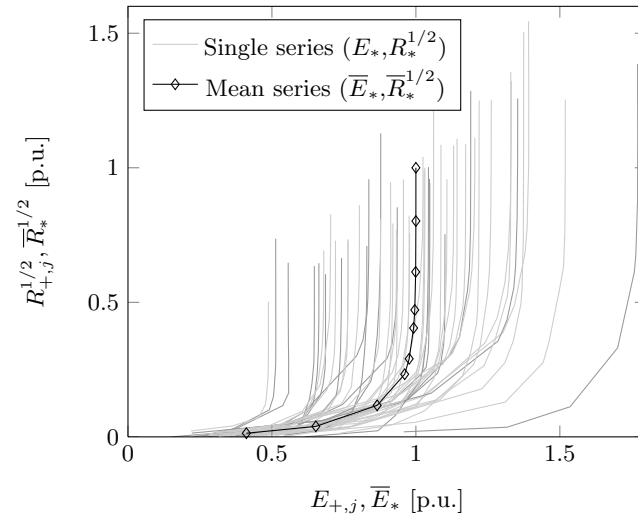
$$\bar{R}_* = \frac{\frac{1}{n} \sum_{j=1}^n \mathcal{J}_{2,j}}{\left[ \frac{1}{n} \sum_{j=1}^n \mathcal{J}_{2,j} \right]^{\alpha=1}}, \quad (16)$$

respectively corresponding to the average energy extraction and average squared variability over  $n$  statistically independent optimization problems using different wind speed series of  $T = 600$  sec that are extracted from the LES data (recall that in total 30000 sec are available for 48 turbines – cfr. §2.2).

In Figure 7 an average over 104 Pareto fronts is shown (summarized into a set  $S_1$  in Table I), together with the first 50 of the Pareto fronts used in the average. To plot these 50 fronts on the same graph, we use

$$E_{+,j} = \frac{\mathcal{J}_{1,j}}{\left[ \frac{1}{n} \sum_{j=1}^n \mathcal{J}_{1,j} \right]^{\alpha=1}}, \quad (17)$$

and a similar definition for  $R_{+,j}$ . The total number  $n$  used for the average in Figure 7 is estimated based on the central limit theorem, requiring a 95% confidence interval on the average  $[(\bar{E}_*, \bar{R}_*^{1/2})]^{\alpha=1} \equiv (1, 1)$  of  $\pm 0.05$ . Thus taking  $\sigma = \max(\sigma_E, \sigma_R)$ , with  $\sigma_E$  and  $\sigma_R$  the standard deviation on  $E_{+,j}$  and  $R_{+,j}^{1/2}$ , we estimate  $n = (2\sigma/0.05)^2 \approx 104$ .



**Figure 7.** Average and individual pareto fronts of each  $T = 600$  s for  $n = 104$  time series from set  $S_1$  in Table I. Only the first 50 time series are shown.

Figure 7 shows that averaging the Pareto front over a sufficient number of independent optimization cases is important to obtain a correct impression of the true potential for power smoothing. A large variance is observed on maximum extracted power and variability  $[E_{+,j}]^{\alpha=1}$ , and  $[R_{+,j}]^{\alpha=1}$ , and next to that, also the slopes of the Pareto fronts depend on the particular 600-sec turbulent-flow realization that is used as an input to the optimization cases. Also the points  $[(0, E_{+,j,min})]^{\alpha=0}$  vary a lot, related to different values of  $P_{min}$  in the time series (cfr. Figure 4b). Thus, without proper averaging, it is difficult to formulate accurate conclusions on the trade-off between power smoothing and energy extraction in optimal control of wind turbines. Therefore, for the remainder of this work, the discussion is based on averaged Pareto fronts using constructed sets summarized in Table I\*.

Set	# Turbines	# Time series	# Pareto points	# Variables simulation	CPU time [h]
$S_1$	1	104	× 10	5600	885
$S_2$	2	60	× 10	11200	1470
$S_{6,\perp}$	6	20	× 10	33600	1145
$S_{6,\parallel}$	6	75	× 10	33600	3050
$S_{24}$	24	17	× 10	134400	2751

**Table I.** Summary of sets used for analysis. Each set has a number of turbines considered in the optimization problem developed, given by the subscript in the set name. Furthermore,  $\perp$  and  $\parallel$  respectively give a set of spanwise and streamwise spaced turbines. The multiplication of the number of time series considered for each set with the number of Pareto points gives the amount of simulations performed for that specific set.

#### 4. COORDINATED CONTROL OF MULTIPLE TURBINES

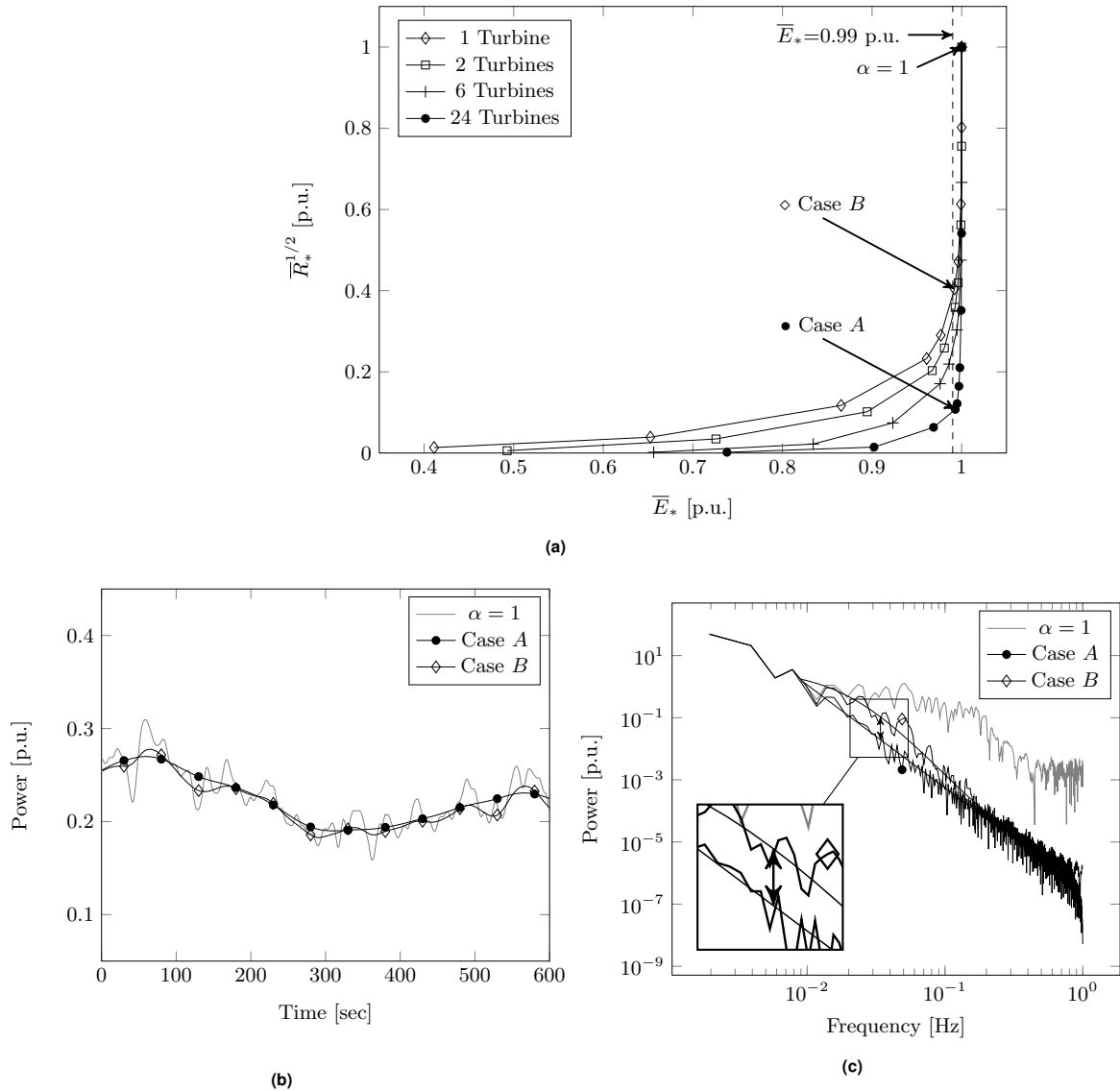
We now investigate the coordinated control of multiple turbines. In §4.1 we discuss the coordination benefit of respectively 2, 6 and 24 turbines in one optimization problem. Secondly, in §4.2, we investigate effects of spatial correlation by comparing the control of a set of turbines along either a row or column of the wind farm.

\*Simulations are performed on double quad-core Xeon 5560 CPU's with 24 GB of shared RAM. Up to 8 nodes, or 64 CPU's, have been used simultaneously on the computing infrastructure of the Flemish Supercomputing Center

#### 4.1. Results for multiple turbines

All previous cases considered only one turbine in the optimization set-up. We now look at the potential benefits of coordinating the optimal control of multiple turbines. Such benefits are analyzed by considering multiple (2, 6 and 24) turbines simultaneously in one optimization set-up, fed by wind series from the LES data set (cfr. §2.2). To obtain the averaged Pareto front,  $n$  different pairs are combined (cfr. sets  $S_2$ ,  $S_{6,\perp}$ ,  $S_{6,\parallel}$  and  $S_{24}$  in Table I) to calculate the averaged values  $\bar{E}_*$  and  $\bar{R}_*^{1/2}$  as function of  $\alpha$  (cfr. §3.3). We keep using the normalization from Eqs. (15), and (16), so that  $[(E_*, R_*^{1/2})]^{\alpha=1} \equiv (1, 1)$ , allowing a direct comparison between the different Pareto fronts in Figure 8a.

For the coordinated cases, the turbines are respectively arranged in a row of 2, a row of 6, and a grid of 6 by 4, respectively for the 2, 6, and 24 turbine case. Obviously, the Pareto frontiers depend on this spatial arrangement and the



**Figure 8.** (a) Average Pareto front for respectively 1, 2, 6 and 24 turbines in a wind farm, representing respectively sets  $S_1$ ,  $S_2$ ,  $S_{6,\perp}$  and  $S_{24}$  in Table I. The power elapse (b) and frequency spectrum (c) are illustrated for one time series handled in two cases: Case A with 24 turbines in the same optimization and Case B with the total power of separate optimization runs for the same 24 turbines. In (b), and (c) results are shown for points on the Pareto frontier for which the energy loss equals 1% compared to control without smoothing ( $\bar{E}_* = 0.99$  p.u.), i.e. case A and B (also indicated on (a)).

spatial correlation of the wind field, but we refer to §4.2 for a detailed discussion on the effects of spatial correlation. Here we focus on a comparison between coordinated and uncoordinated turbines.

As point of comparison for, e.g., the 24 coordinated turbine case, we need the Pareto front corresponding to 24 uncoordinated turbines subject to the same wind field. However, it is easily shown that the average over 24 uncoordinated optimally controlled turbines converges to the same normalized Pareto front as that of the averaged one-turbine case independent of spatial correlation (remark that the total level of variability before normalization does depend on spatial correlation). The same argument holds for the 2 and 6 turbine cases in Figure 8a. Thus, we can compare all correlated Pareto fronts to the one-turbine Pareto front to identify the possible gains of coordination.

The comparison of all four Pareto fronts in Figure 8 demonstrates a clear benefit of using coordinated control. The Pareto front for the coordinated control case is steeper, and extracts more energy for the same level of variability. Also the minimum power at zero variability,  $\bar{E}_{*,min}$ , is higher for the coordinated control case. Such benefits of coordinated control become larger when more turbines are added. The higher the number of turbines, the more variability can be reduced without losing significant amounts of extracted energy. In a set of turbines, turbulent wind-speed fluctuations partially compensate each other, i.e. the wind accelerates at some turbines, but decelerates at others (cfr. e.g., De Rijcke *et al* [26] for a detailed discussion on a two-turbine case). Thus, the amount of rotating kinetic energy reserves increases relative to the accumulated variability, so that optimal control of power smoothing becomes more effective.

We further compare two cases A and B, where Case A corresponds to the coordinated optimal control of 24 turbines, and Case B the same 24 turbines that are individually controlled. Both cases are selected thus on the respective Pareto frontiers that their energy loss is 1% compared to optimal control of energy extraction at  $\alpha = 1$ . This is also marked on Figure 8a. In Figure 8b the farm-power is shown for one time series in Case A, and Case B, showing that the former leads to a smoother signal in time. In Figure 8c, we further compare Case A and Case B in terms of their averaged power spectral density. We observe that both Cases are efficient in reducing variability for time scales below 10 sec ( $> 0.1$  Hz). However, when looking at time scales ranging from 10 to 100 sec, the case with 24 coordinated turbines (Case A) is much more efficient in reducing variability. For a time scale of 50 sec (0.02 Hz) Case A reduces variability by a factor 6, compared to 1.4 for Case B, i.e. Case A is a factor 4 better. Around 20 sec (0.05 Hz) we find that Case A is a factor 5 better than Case B.

## 4.2. Impact of spatial correlation

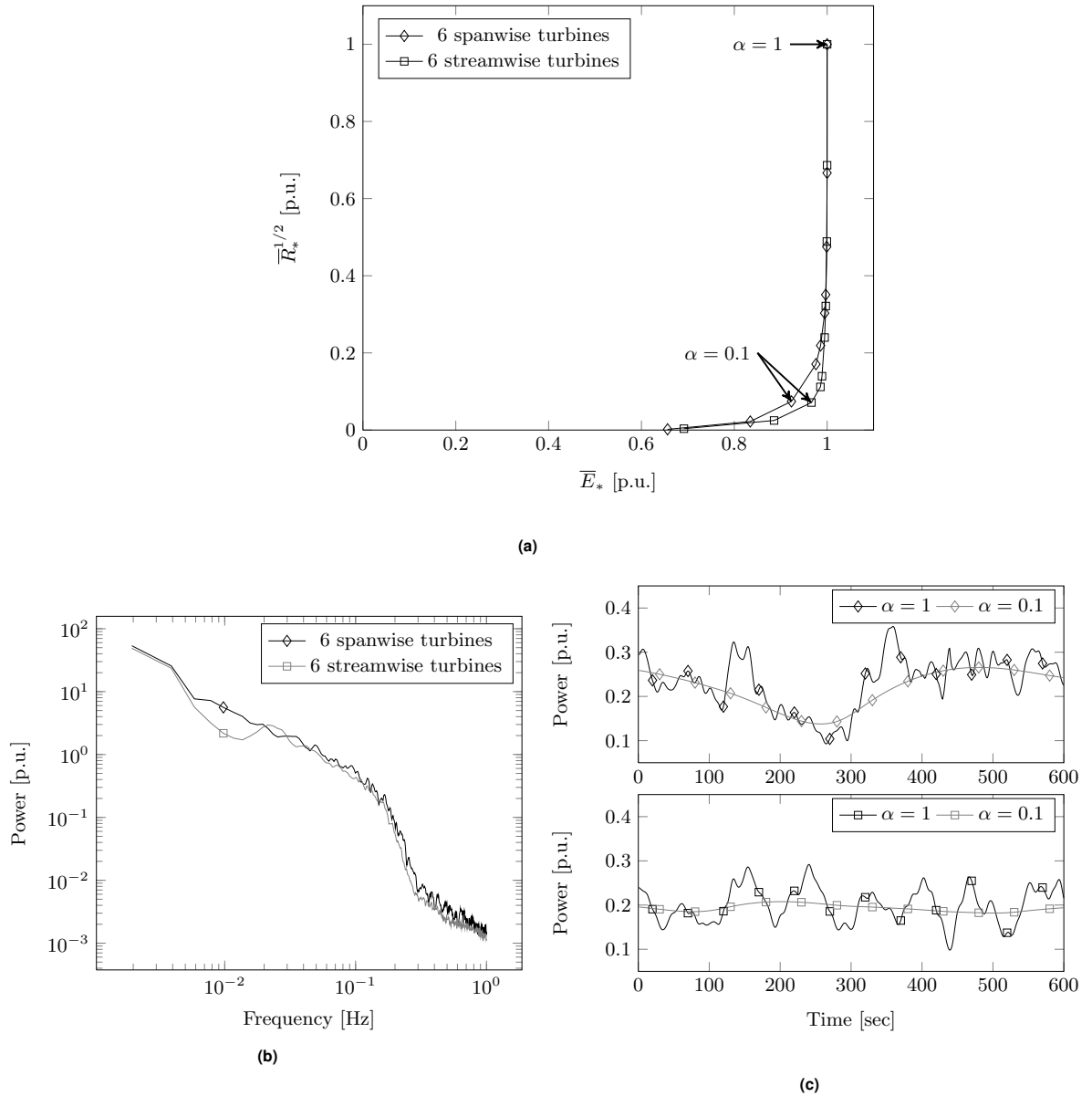
Spatial placement of turbines and the associated spatial correlations in the turbulent flow are expected to influence the character of power variations and consequently the extent to which power can be smoothed by optimal control. In the current section, we investigate possible effects of correlation by comparing the coordinated control of six turbines that are either located in a row parallel to the prevailing wind (streamwise), or in a row that is perpendicular to the prevailing wind (spanwise). For both cases, again a sufficient number of time series is combined into set  $S_{6,\perp}$  and  $S_{6,\parallel}$  (cfr. Table I), respectively for the spanwise and streamwise case.

Figure 9a shows the difference in Pareto frontiers between both cases. Smoothing spanwise spaced turbines is clearly more difficult than streamwise turbines, and infers a higher cost for the same reduction of variability. On top of the larger relative reduction in variability, the streamwise case has a lower absolute variability to start with. This is illustrated in Figure 9b, where the power spectra for both cases are shown at  $\alpha = 1$ . In particular for frequencies around approximately 10 mHz, a large difference in spectral density is observed. This frequency roughly corresponds with the flow convection time between consecutive turbines in the streamwise case, i.e. a turbine spacing of  $s_x D = 785$  m, and an average hub-height velocity of 7.5 m/s (cfr. Figure 2a) yields approximately 9 mHz. In fact, in the streamwise case, large turbulent flow structures with streamwise size proportional to the turbine spacing, e.g., a wind gust, can lead to an increase in wind speed at one of the turbines, while effecting a decrease of wind speed at a neighbour turbine, naturally decreasing variability.

Finally, in Figure 9c, time series of the farm power are shown for the two different cases, both for  $\alpha = 1$ , and  $\alpha = 0.1$ . It is appreciated that the streamwise case is much smoother at  $\alpha = 0.1$ , while keeping the energy loss limited to only 3.4%, while this is 7.7% for the spanwise case (cfr. Figure 9a). Note that also an energy loss of 3.4% may be unacceptable, and points higher up the Pareto front (at lower values of energy loss) may need to be selected.

## 5. CONCLUSIONS

In the current study, we investigated optimal control of rotating kinetic energy reserves in large wind farms, with the aim to smooth power variability originating from turbulent fluctuations of wind speeds. Turbulent wind data were obtained from large-eddy simulations of a wind farm boundary layer. Pareto frontiers were constructed, showing the optimal trade-off between reducing power variability, and maximizing energy extraction in the wind farm. These Pareto frontiers were averaged over a large number of optimal-control realizations, reducing the statistical error to 5% (using a 95% confidence interval). Different cases were compared, i.e., uncoordinated control of a set of turbines, and coordinated control in groups of 2, 6, and 24 turbines.



**Figure 9.** (a) Pareto fronts for six turbines respectively arranged perpendicular (spanwise) to the prevailing wind (set  $S_{6,\perp}$ ), and parallel (streamwise) with the prevailing wind (set  $S_{6,\parallel}$ ). (b) Average power density spectrum for  $\alpha = 1$ . (c) Power elapse for one time series from both sets in (a) when setting  $\alpha$  respectively to 1 and 0.1, where the upper and lower graph are respectively for a set of spanwise and streamwise spaced turbines.

We observed that coordinated control allows to smooth power considerably more effective than similar cases where turbines are controlled individually. In particular variability over longer time scales (10–100 sec) is reduced considerably more for coordinated control. For instance, restricting the energy-loss incurred by smoothing to 1%, and looking at frequencies around 20 mHz (i.e. about 50 sec), we manage to reduce variability with a factor of 6 for a coordinated case with 24 turbines, compared to a factor of 1.4 only for an uncoordinated case. We further also observed that spatial coherence of the turbulent flow field is important for power smoothing. Turbines that are aligned in a streamwise row allow for easier smoothing than turbines in a cross-stream row.

The current study focussed on optimal control of power smoothing in wind farms, and the development of a methodology to compare possible gains of control approaches in a variable turbulent wind field. This methodology can also be used to assess the gain of smoothing power variations within specified frequency bandwidths (e.g. by using different definitions for the cost function  $J_2$  weighted over a range of time lags  $\tau$ ), or when turbines are subject to rotor speeds limits.

Furthermore, we believe that the proposed framework yields an important benchmark for future development of real controllers. It also provides insights on the value of short-term forecasts and coordination among turbines, two aspects that are crucial for the development of this approach into wind turbine and wind farm controllers. To properly exploit the value of forecasts and coordination in real controllers, the integration of online measurements (e.g. by lidars) and communication among turbines are necessary. This is currently subject of further research.

## ACKNOWLEDGEMENT

The authors acknowledge financial support from the Research Foundation – Flanders (FWO, grant no. G.0376.12). J.M. further acknowledges support from the European Research Council (FP7–Ideas, grant no. 306471). Simulations were partly performed on the computing infrastructure of the VSC Flemish Supercomputer. The authors further thank TOMLAB for support on their optimization software.

## APPENDIX I: OPTIMALITY CONDITIONS AND FORMULATION OF COST-FUNCTION GRADIENTS

The minimum principle allows to calculate the minimum of the functional  $\mathcal{J}$  defined in equation (4) constrained by the state equation of each turbine [27]. By constructing the Lagrangian  $L$ , this principle leads to the equations that are the core of the algorithm. This Lagrangian adjoins the constraints on the system dynamics to the objective  $J$  by introducing time-varying Lagrange multipliers  $\omega^*$ , resulting in the functional  $\mathcal{L}$ :

$$\begin{aligned}\mathcal{L}(\omega, \beta, T_e, \omega^*, T) &= \int_0^T L(\omega, \beta, T_e, \omega^*, t) dt \\ &= \int_0^T \left[ J(\omega, T_e, t) + \sum_{i=1}^N \left( J \frac{d\omega}{dt} - T_{a,i} + T_{e,i} \right) \omega_i^* \right] dt.\end{aligned}\quad (18)$$

The minimum principle states that the optimal trajectory for  $\omega$ ,  $\beta$ ,  $T_e$  and the corresponding Lagrange multipliers  $\omega^*$ , must minimize the Lagrangian. Thus, the necessary conditions for obtaining an extremum are derived by the method calculus of variations, i.e. the derivatives of the functional  $\mathcal{L}$  with respect to all variables should be zero, or

$$\frac{\partial \mathcal{L}}{\partial \omega_i} = 0, \quad \frac{\partial \mathcal{L}}{\partial \beta_i} = 0, \quad \frac{\partial \mathcal{L}}{\partial T_{e,i}} = 0, \quad \frac{\partial \mathcal{L}}{\partial \omega_i^*} = 0 \quad \text{for } i = 1 \dots N \quad (19)$$

Working out these derivatives gives

$$\frac{\partial \mathcal{L}}{\partial \omega_i^*} \delta \omega_i^* = 0 + \int_0^T \left[ I \frac{d\omega_i}{dt} - T_{a,i} + T_{e,i} \right] dt \quad (20)$$

$$\frac{\partial \mathcal{L}}{\partial \omega_i} \delta \omega_i = \int_0^T \frac{\partial J}{\partial \omega_i} \delta \omega_i dt + \int_0^T \left[ I \frac{d\delta \omega_i}{dt} - \frac{\partial T_{a,i}}{\partial \omega_i} \right] \omega_i^* dt \quad (21)$$

$$\frac{\partial \mathcal{L}}{\partial \beta_i} \delta \beta_i = \int_0^T \frac{\partial J}{\partial \beta_i} \delta \beta_i dt - \int_0^T \left[ \frac{\partial T_{a,i}}{\partial \beta_i} \right] \delta \beta_i dt \quad (22)$$

$$\frac{\partial \mathcal{L}}{\partial T_{e,i}} \delta T_{e,i} = \int_0^T \frac{\partial J}{\partial T_{e,i}} \delta T_{e,i} dt + \int_0^T \omega_i^* \delta T_{e,i} dt \quad (23)$$

Setting the first derivative from equation (20) to zero results in the state equations, earlier stated as a constraint in Eq. (24):

$$J \frac{d\omega_i}{dt} = T_{a,i} - T_{e,i} \quad \text{for } i = 1 \dots N. \quad (24)$$

By performing the principle of integrating by parts on the derivative from (21) and setting it to zero, the second condition becomes

$$\int_0^T \frac{\partial J}{\partial \omega_i} \delta \omega_i dt + \int_0^T \left[ I \frac{d\omega_i^*}{dt} \delta \omega_i - \frac{\partial T_{a,i}}{\partial \omega_i} \omega_i^* \delta \omega_i \right] dt + J \omega_i^* \delta \omega_i \Big|_0^T = 0 \quad (25)$$

The last part in this equation gives the boundary conditions. At time  $t = 0$ ,  $\delta \omega_i$  is chosen to be zero. In addition,  $\omega^*(T)$  is set to 0, which is the transversality condition for the free-terminal-state problem, i.e. when the state variable  $\omega$  at time  $T$

is free [27]. With these additional conditions, the adjoint equations are defined for every  $\delta\omega_i$

$$\frac{\partial J}{\partial \omega_i} = -J \frac{d\omega_i^*}{dt} + \frac{\partial T_{a,i}}{\partial \omega_i} \omega_i^* \quad \text{for } i = 1 \dots N. \quad (26)$$

$$\frac{d\omega_i^*}{dt} = \frac{1}{J} \left[ \frac{\partial T_{a,i}}{\partial \omega_i} \omega_i^* - \frac{\partial J}{\partial \omega_i} \right] \quad \text{for } i = 1 \dots N. \quad (27)$$

In addition, equations (22) and (23) give a weak representation of the minimum principle. Consequently, for every  $\delta\beta_i$  and  $\delta T_{e,i}$ , the following equations should be valid to obtain a local minimum:

$$-\frac{\partial T_{a,i}}{\partial \omega_i} \omega_i^* + \frac{\partial J}{\partial \beta_i} = 0 \quad \text{for } i = 1 \dots N, \quad (28)$$

$$\omega_i^* + \frac{\partial J}{\partial T_{e,i}} = 0 \quad \text{for } i = 1 \dots N.. \quad (29)$$

The set of equations (24), (27), (28) and (29) provides the necessary conditions for an optimal trajectory. Moreover, these equations are the core of the algorithm for a gradient-based solver. First the state equation (24) is solved forward in time. The states are used to solve the adjoint equation (27) backwards in time to obtain the Lagrange multipliers  $\omega_i^*$ . The last two equations give the derivative of the Lagrangian  $\mathcal{L}$  to both control variables  $\beta$  and  $T_e$ , i.e. the gradient  $\nabla_{\beta, T_e} \mathcal{L}$ . This approach to calculate the gradient, i.e. the adjoint method, is most suited to calculate the gradient of a scalar function with respect to a high number of parameters [23]. Contrary to the forward method or using finite differences, this method's computational effort is relatively insensitive to the number of parameters. This argument is highly relevant in view of using the gradient calculation as input for a gradient based solver.

## REFERENCES

1. European Wind Energy Association. The european offshore wind industry - key trends and statistics 2011. *Technical Report*, EWEA 2011.
2. European Network of Transmission System Operators for Electricity. Draft network code on load-frequency control and reserves. *Technical Report*, ENTSO-E 2013.
3. Eirgrid. Eirgrid grid code version 3.3: Wind grid code only. *Technical Report*, Eirgrid 2009.
4. Energinet. Technical regulation 3.2.5 for wind power plants with a power output greater than 11 kW. *Technical Report*, Energinet 2010.
5. Albadi M, El-Saadany E. Overview of wind power intermittency impacts on power systems. *Electric Power Systems Research* 2010; **80**(6):627 – 632.
6. Lebron J, Castillo L, Meneveau C. Experimental study of the kinetic energy budget in a wind turbine stream-tube. *Journal of Turbulence* 2012; **13**(23).
7. Nanahara T, Asari M, Sato T, Yamaguchi K, Shibata M, Maejima T. Smoothing effects of distributed wind turbines. part 1. coherence and smoothing effects at a wind farm. *Wind Energy* 2004; **7**(2):61–74.
8. Nanahara T, Asari M, Maejima T, Sato T, Yamaguchi K, Shibata M. Smoothing effects of distributed wind turbines. part 2. coherence among power output of distant wind turbines. *Wind Energy* 2004; **7**(2):75–85.
9. Sørensen P, Cutululis N, Viguera-Rodríguez A, Jensen L, Hjerrild J, Donovan M, Madsen H. Power fluctuations from large wind farms. *Power Systems, IEEE Transactions on* 2007; **22**(3):958–965.
10. Sørensen P, Cutululis NA, Viguera-Rodríguez A, Madsen H, Pinson P, Jensen LE, Hjerrild J, Donovan M. Modelling of power fluctuations from large offshore wind farms. *Wind Energy* 2008; **11**(1):29–43.
11. Ran L, Bumby J, Tavner P. Use of turbine inertia for power smoothing of wind turbines with a DFIG. *Harmonics and Quality of Power, 2004. 11th International Conference on*, 2004; 106–111.
12. Changling L, Banakar H, Shen B, Boon-Teck Ooi. Strategies to smooth wind power fluctuations of wind turbine generator. *Energy Conversion, IEEE Transactions on* 2007; **22**(2):341–349.
13. Senjyu T, Sakamoto R, Urasaki N, Funabashi T, Fujita H, Sekine H. Output power leveling of wind turbine generator for all operating regions by pitch angle control. *Energy Conversion, IEEE Transactions on* 2006; **21**(2):467–475.
14. Lubosny Z, Bialek J. Supervisory control of a wind farm. *Power Systems, IEEE Transactions on* 2007; **22**(3):985–994.
15. Clemow P, Green T, Hernandez-Aramburo C. Wind farm output smoothing through co-ordinated control and short-term wind speed prediction. *Power and Energy Society General Meeting, 2010 IEEE*, 2010; 1–8.
16. Calaf M, Meneveau C, Meyers J. Large eddy simulations study of fully developed wind-turbine array boundary layers. *Physics of fluids* 2010; **22**:015 110.
17. Meyers J, Meneveau C. Flow visualization using momentum and energy transport tubes and applications to turbulent flow in wind farms. *Journal of Fluid Mechanics* 2013; **715**:335–358.



18. Burton T. *Wind Energy Handbook*. Wiley & Sons, 2001.
19. Goit J, Meyers J. Optimal control of wind farm power extraction in large eddy simulations. *AIAA SciTech 32nd ASME Wind Energy Symposium*, American Institute of Aeronautics and Astronautics, 2014.
20. Slootweg JG, Polinder H, Kling WL. Dynamic modelling of a wind turbine with doubly-fed induction generator. *Vancouver IEEE Power Engineering Society Summer Meeting*, 2001.
21. Heier S. *Grid integration of wind energy conversion systems*. Wiley, 2005.
22. Rutquist P, Edvall M. *User's Manual for TOMLAB*. Tomlab Optimization Inc.: 1260 SE Bishop Blvd Ste E, Pullman, WA 99163, USA, 2008.
23. Cao Y, Li S, Petzold L, Serban R. Adjoint sensitivity analysis for differential-algebraic equations: the adjoint dae system and its numerical solution. *SIAM journal on scientific computing* 2003; **24**:1076–1089.
24. Gros S. An economic nmpe formulation for wind turbine control. *52nd IEEE Conference on Decision and Control*, 2013; Paper no. TuB11.4.
25. Marler R, Arora J. The weighted sum method for multi-objective optimization: new insights. *Structural and Multidisciplinary Optimization* 2010; **41**(6):853–862.
26. De Rijcke S, Meyers J, Driesen J. Reducing power gradients in large-scale wind farms by optimal active power control. *IEEE PowerTech Grenoble*, 2013.
27. Chiang. *Elements of Dynamic Optimization*. McGraw-Hill, 1992.